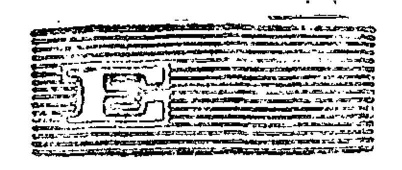


NACIONES UNIDAS

CONSEJO ECONOMICO Y SOCIAL



LIMITADO

ST/ECLA/Conf.43/L.3
17 de julio de 1972

ESPAÑOL
ORIGINAL: INGLES

COMISION ECONOMICA PARA AMERICA LATINA

SEMINARIO SOBRE LA PREPARACION Y USO DE TABULACIONES DE CENSOS DE POBLACION Y HABITACION

Organizado por las Naciones Unidas, por intermedio de la Comisión Económica para América Latina, la Oficina de Estadística de las Naciones Unidas, el Fondo de las Naciones Unidas para Actividades en materia de Población con la colaboración del Centro Latinoamericano de Demografía

Santiago de Chile, 14 al 19 de agosto de 1972

CENTS, UNA TECNICA PARA LA TABULACION CON COMPUTADORA DE LOS RESULTADOS DE CENSOS DE POBLACION Y HABITACION

por Howard G. Brunsman
de la Oficina del Censo de los Estados Unidos

72-7-1752

En los últimos años se ha ido desarrollando un nuevo método de tabulación de los datos obtenidos de los censos de población y habitación. El sistema CENTS es una variante del método básico empleado por la Oficina del Censo de los Estados Unidos para el procesamiento de sus últimos censos. El sistema comprende, esquemáticamente, los pasos siguientes:

1) Determinar las zonas para las cuales se requiere un grupo de tabulaciones.

2) Agrupar los registros básicos de manera que ningún grupo contenga registros de más de una zona para la cual se requieren tabulaciones. Por ejemplo, algunas tabulaciones de los resultados censales de los Estados Unidos se especifican para los lugares urbanos, las zonas urbanizadas, las zonas metropolitanas y los condados, Estados, divisiones y regiones, así como las partes urbanas y rurales de tales subdivisiones. Para preparar tales tabulaciones los registros deben agruparse por cada lugar urbano, el resto de registros urbanos y los registros rurales para cada condado. Si una zona urbana se encuentra en dos o más condados, se agruparán separadamente. De igual manera cuando un lugar urbanizado se encuentre en dos o más condados, las partes del restante de una zona urbanizada serán agrupadas separadamente.

3) Preparar la tabulación básica para cada uno de estos subgrupos.

4) Consolidar las tabulaciones según las zonas que se emplearán para la publicación.

5) Preparar las tablas que se publicarán a partir de las cifras consolidadas.

En el ejemplo que se empleó, la tabulación se preparó para unos 9 000 grupos. Las tabulaciones de estos grupos se consolidaron en cifras para 5 445 lugares urbanos distintos y 213 zonas urbanizadas. Se hizo lo mismo para el total y las partes urbanas y rurales de 3 134 condados, 50 Estados, 9 divisiones, 4 regiones y el total de los Estados Unidos. Asimismo, se consolidaron las tabulaciones en cifras para 212 zonas estadísticas metropolitanas uniformes, y para 509 zonas económicas estatales, que están formadas por agrupaciones de condados.

/En el

En el censo de 1950 estas operaciones se efectuaron empleando el registro básico en tarjetas perforadas. Las tabulaciones fueron preparadas por equipos unitarios (contadores). Los datos resultantes se traspasaron a tarjetas perforadas sumarias que se clasificaron y procesaron con tabuladoras convencionales para consolidar los datos; luego éstos se incluyeron en cuadros que debieron mecanografiarse para su publicación. En los censos de 1960 y 1970 las mismas operaciones básicas se cumplieron con la computadora.

La Oficina del Censo de los Estados Unidos tiene dos ventajas fundamentales sobre los países más pequeños para la realización de estas operaciones. En primer lugar, tiene una computadora grande. En segundo lugar, cuenta con un grupo numeroso de personal especializado en preparar los programas. Desafortunadamente pocos son los países en desarrollo que cuentan con estas ventajas. Debiera ser posible aprovechar los conjuntos de programas (software packages) disponibles para reducir los esfuerzos de programación. Hay muchos disponibles. Entre ellos RPG (Programa Generador de Informes), DATATEXT, 360 CROSSTAB Y MINITAB. Todos ellos tienen sus respectivas limitaciones. Algunos requieren una computadora con una capacidad mayor que la disponible en muchos países en desarrollo. Otros son ineficientes por el pequeño número de tabulaciones que pueden producir en una sola operación. Y otros son incapaces de consolidar los datos tabulados para todos los temas de publicación que interesan con relación a un censo.

Al reconocer la necesidad de un conjunto de programas que pudiera procesar eficientemente una gran cantidad de datos en una computadora relativamente pequeña, desarrollamos la técnica de programación, o software package que se ha llegado a conocer con el nombre de CENTS (Census Tabulation System). El sistema fue desarrollado principalmente por la Oficina del Censo de los Estados Unidos, bajo el auspicio de la Agencia para el Desarrollo Internacional del Gobierno de los Estados Unidos, pero debe reconocerse que varios gobiernos de este hemisferio, así como la Organización de los Estados Americanos, apoyaron el proyecto y que el sistema está en uso en algunos de los países que colaboraron.

El método CENTS ha sido empleado para tabular los datos de habitación y población de un total de 50 000 000 habitantes de México. Este censo se levantó en enero de 1970 y muchas de sus tabulaciones básicas estaban listas en noviembre del mismo año. Se empleó en Chile para tabular por adelantado datos muestrales del censo de 1970 y en Panamá para tabular los censos de vivienda y población de 1970. Se encuentra en aplicación por lo menos en otros veinte países de todo el mundo (Grecia, Vietnam, Kenia, Brasil, Costa Rica, Filipinas, Jamaica y Guatemala, entre otros).

/El CENTS

El CENTS ha sido especialmente preparado para procesar el volumen más grande posible de trabajo en el tiempo mínimo que permite la capacidad de la computadora. Todas las operaciones básicas de tabulación se efectúan mediante programas especiales de computación activados por tarjetas-parámetro. La tabulación básica que supone el paso 3 anterior, es realizada por el programa CENTAL activado por 4 a 8 tarjetas-parámetro para cada cuadro. La agrupación secuencial y la consolidación de los datos acumulados (paso 4) se lleva a cabo con las rutinas estandarizadas de agrupación secuencial del programa CENCON, que son activadas por 4 a 10 tarjetas-parámetro por juego de cuadros. La preparación de los cuadros para la publicación (paso 5) se realiza con el programa CENPREP. Este programa requiere un tarjeta con nombre de zona para cada zona. Cada cuadro exige de 4 a 10 tarjetas-parámetro para los encabezamientos y los cálculos, más una tarjeta para el título de cada línea a la izquierda del cuadro. Se facilita la preparación de las tarjetas parámetro del CENTS con el empleo de formularios especiales de codificación. Las tarjetas-parámetro de CENTAL se codifican en el Formulario PH-421, según aparece en la muestra 1. Las tarjetas-parámetro de CENCON se codifican en el Formulario PH-422, según aparece en la muestra 2. Las diversas tarjetas-parámetro para el CENPREP se codifican en los Formularios PH-423 a PH-426 (véanse las muestras 1 a 4) 1/.

Supongamos, por ejemplo, que queremos tabular el cuadro que figura en la muestra 1: población por relación con jefe del hogar por edad. Suponemos que no se necesita la tabulación para subdivisiones del universo sino para el total de todos los registros. El código de la edad según aparece en el byte 15 de cada registro de población es 00 para los menores de un año, 01 a 98 de 1 a 98 años, y 99 para 99 años y más. Los códigos de relación en el byte 12 son los siguientes:

Códigos de relación

Relación	Código S15
Jefe de hogar	3
Esposa o conviviente	4
Hijo del jefe	5
Otro pariente del jefe	6
Sin parentesco con el jefe	7
Vive solo	8

1/ Estas muestras aparecen en el documento informativo: U.S. Bureau of the Census, CENTS, An approach to a computer tabulation of data from a population and housing census, Series ISP 4, Nº 1, Washington D.C., 1970, Appendix II.

/Debemos informar

Debemos informar a la computadora que el cuadro 1 contiene 11 columnas y 7 líneas; que la persona debe incluirse en la columna 2 si su código de edad es 4 o menos; en caso contrario, en la columna 3 si es de 9 o menos; de no ser así, en la columna 4 si el código de edad es 14 o menos, etc., e incluirlo en la columna 11 si el código de edad es 99 o menos.

También debemos instruir a la computadora que cuente en la línea 2 si el código de relación es 3 o menos; de no ser así, en la 4 si es 4 o menos; en la 5 si es 5 o menos, etc.

Con el sistema CENTS, estas operaciones se realizan al suministrar a la computadora las tarjetas-parámetro que figuran en la muestra 1. La primera tarjeta dimensiona el cuadro; la segunda instruye a la computadora que proceda a la recodificación de la relación sobre la base del byte 12; la tercera le ordena establecer las recodificaciones de edad, a base del byte 15. La cuarta, tabular en el cuadro 1 en que la columna ha sido especificada por la recodificación de edad y la línea por la recodificación de relación. Esta parte del programa CENTS produce matrices de valores tabuladas hasta ese momento para el cuadro. Debemos ahora derivar los totales de las diversas columnas y líneas y presentar los datos en forma legible, con encabezamientos apropiados y rotulado de las líneas. Estas especificaciones del cuadro se las dan a la computadora en las tarjetas-parámetro que aparecen en la muestra 1. La primera tarjeta especifica que el espacio para título a la izquierda del cuadro tiene 33 posiciones y que los datos se presentan en 11 columnas del tipo 1, cada una de 9 columnas de ancho. La próxima línea especifica que por cada línea del cuadro deseamos calcular la columna 1 como la suma de las columnas 2 a 11. En la próxima línea se dice que para cada columna del cuadro deseamos calcular la línea 1 como la suma de las líneas 2 a 7. Las próximas 6 líneas muestran las partes izquierda y derecha de cada una de las 3 líneas que encabezan el cuadro. Las próximas 7 líneas dan el contenido de los títulos que figuran al lado izquierdo para las 7 líneas del cuadro.

Evidentemente, se ha presentado un ejemplo relativamente sencillo. Hay más instrucciones en las operaciones de tabulación. Hay otros dos tipos de recodificación en el CENTAL para satisfacer otras condiciones. Con otras instrucciones pueden omitirse ciertos casos del cuadro, como las personas menores de 12 años. Un cuadro puede contener un grupo especial de líneas para los varones y otro para las mujeres. Una instrucción permitirá aumentar el código por línea en una cantidad determinada cuando el sexo es femenino.

/Otras instrucciones

Otras instrucciones del CENPREP permitirán preparar cifras derivadas, como el tamaño medio del hogar, distribuciones porcentuales, y medianas. Cada una de estas cifras derivadas se obtiene con una sola tarjeta-parámetro.

El ejemplo presentado anteriormente supone que la tabulación se necesita sólo para el total de todos los casos. Cuando se necesitan datos por separado para diversas subdivisiones geográficas, es preciso incluir entre las tarjetas-parámetro del CENTAL, una tarjeta con especificación de zona que designe las zonas más detalladas para las cuales se requieren los datos; agrupar secuencialmente la salida del CENTAL con una rutina estándar de secuencia; consolidar los datos en las zonas que se van a publicar usando el programa CENCON activado por tarjetas-parámetro; agrupar secuencialmente la tabulación por lugares en el orden especificado para la publicación y especificar en CENPREP las zonas para las cuales se imprimirá la tabulación. Las tarjetas-parámetro necesarias para llevar a cabo estas operaciones aparecen en la muestra 2. Las tarjetas-parámetro requeridas para la tabulación de otros cuadros aparecen en las muestras 3 y 4.

El uso del CENTS en la tabulación de un censo presenta las ventajas siguientes:

1) Una enorme reducción en horas-hombre de programación. Toda la programación de un cuadro típico no demora más que unas pocas horas.

2) Se requieren menos programadores capacitados. Parte del personal de programación debe tener un conocimiento completo de la técnica de computación, pero la mayoría de las tarjetas-parámetro pueden ser preparadas por personas menos calificadas. Las tarjetas de especificación de impresión pueden ser preparadas por personas sin experiencia anterior alguna en programación.

3) Se requiere menos tiempo de operación de la computadora. Esta ventaja resulta de los factores siguientes:

a) El programa es por lo menos tan rápido como los programas normales escritos en lenguaje ensamblador, ya que también está escrito en ese lenguaje. En un ensayo efectuado en México, se tabularon 12 cuadro, a razón de 11 000 personas por minuto, en una computadora 360/40 y 27 000 por minuto en una 360/50. En el Brasil se tabularon 18 cuadros a razón de 2 000 personas por segundo en una 370/155.

/b) Las

b) Las tarjetas-parámetro se procesan en el "tiempo objeto". Esto ahorra de 10 a 30 minutos del tiempo de operación de la computadora requerido para ensamblar cada nuevo programa con las técnicas de programación corrientes.

c) Una parte relativamente grande de la memoria de la computadora se encuentra disponible para almacenar los datos tabulados con el programa CENTAL. Ello deriva de la pequeña porción de memoria interna que requiere el CENTS en comparación con FORTRAN, COBOL y otros programas. La porción del programa CENTAL que convierte el contenido de las tarjetas-parámetro en instrucciones de operación, comparte las posiciones de la memoria con los datos tabulados. Una computadora con sólo 32 000 bytes de memoria interna, puede procesar hasta 20 cuadros con un total de 3 500 celdas de conteo.

d) El programa CENTAL permite la tabulación del máximo de cuadros en cada operación, al ajustarse automáticamente a la naturaleza de las tabulaciones y a la cantidad de posiciones de memoria asignadas a la operación.

4) Facilidad para repasar las especificaciones con el analista. El contenido de las tarjetas-parámetro es relativamente simple. El programador puede repasar las especificaciones con el analista y confirmar que se están cumpliendo esas especificaciones.

¿Cuáles son las limitaciones del métodos CENTS? Yo estimo que son las cuatro siguientes:

1) Como fue preparado para computadoras pequeñas, se escribió en un lenguaje básico de máquina y no en FORTRAN O COBOL. Como se aplicó primero en IBM 360, se escribió en un lenguaje ensamblador básico para estas computadoras. Pueden usarse los programas en las computadoras Siemens, que son plenamente compatibles con la IBM 360, y se han adaptado a la UNIVAC 9400, habiendo sido empleados en esa computadora en Grecia. Estamos terminando algunos experimentos para rescribir el programa en COBOL, de modo que pueda aplicarse en una variedad mayor de computadoras.

2) La documentación del CENTS no incluye materiales didácticos. Entre los materiales figuran un juego de hojas de codificación que facilitan la preparación de las tarjetas-parámetro y el listado posterior de los diversos programas de computación. También incluyen la publicación

/"XWN CENTS,

"XWN CENTS, an approach to the Computer tabulation of data from a population and housing census" 2/, que contiene la descripción de cada tarjeta-parámetro y describe cómo prepararla y usarla.

3) Aunque el material CENTS abarca el proceso de tabulación, se supone que el registro básico ha sido revisado y que se han subsanado las omisiones e incongruencias. Lo que se necesita es un sistema de procesamiento que vaya desde la tarjeta perforada hasta el cuadro publicado. La tabulación con los medios tradicionales es demorosa y demanda mucho trabajo. Pero la preparación de un buen programa de corrección exige un programador más calificado que el que requiere el programa de tabulación. Por fortuna, la corrección sólo se lleva a cabo una sola vez, y generalmente puede hacerse con un solo programa de computación, mientras que la tabulación de numerosos cuadros requiere muchos programas.

Después de haber desarrollado los programas CENTS, nos vimos abocados al problema de cómo era mejor presentarlos a sus posibles usuarios. Decidimos organizar seminarios de trabajo. Bajo los auspicios de la Agencia para el Desarrollo Internacional de los Estados Unidos, hemos llevado a cabo muchos de esos seminarios: 6 en inglés en Washington, y 2 en español, uno en Santiago de Chile en septiembre-octubre de 1970 y otro en San José de Costa Rica, en mayo-junio de 1971. Se proyecta realizar otros dos en Africa este año, uno en inglés y el otro en español. Cada seminario dura aproximadamente unas 8 semanas, y han asistido a ellos unos 80 participantes. Los participantes son programadores de los países en desarrollo, cuya asistencia es patrocinada por la AID (aunque 21 de ellos fueron becados por las Naciones Unidas). El curso de instrucción los capacita para usar el CENTS en la tabulación de un censo. También se da alguna instrucción en cómo funcionan los programas CENTS, pues se reconoce que puede haber problemas en hacer funcionar los programas cuando se establecen. También puede ser necesario efectuar tabulaciones que no existen en los programas actuales. El programa de capacitación comprende la posibilidad de una consulta posterior con un representante de la Oficina del Censo de los Estados Unidos quien prestará su asistencia para hacer que el sistema funcione bien.

2/ *Op.cit.*, véase la nota 1/.

/Los materiales

Los materiales CENTS actuales están centrados en el procesamiento de un censo nacional de población y habitación. La entrada puede consistir enteramente de registros de habitación o enteramente de registros de población, o puede estar compuesta de un registro de habitación seguido por un registro separado para cada persona de la unidad de habitación. El sistema permite el procesamiento de los datos mencionados en su totalidad, en que cada registro tiene una ponderación de 1 o de datos de una muestra censal, en que cada registro tiene una ponderación que se muestra en una parte especial del registro. En una versión anterior, el programa CENTS exigía que la entrada se presentara en números binarios de uno o dos bytes y que los registros estuvieran en cinta. Se ha modificado ahora para que pueda aceptar de uno a cuatro dígitos en cinta o en tarjetas perforadas.

¿Cómo se aplica el CENTS a otros problemas de tabulación? Tiene plena aplicación en la tabulación de los resultados de encuestas por muestreo. La facilidad y bajo costo de preparar tabulaciones con CENTS es especialmente ventajosa en las encuestas por muestreo porque a menudo hay escasos fondos disponibles para la tabulación. El programa CENTS se ha empleado también para tabular los censos de agricultura de México, Panamá y Nicaragua. El programa no puede establecer recodificaciones para variables de entrada superiores a 9 999, pero puede verificar si la variable tiene un valor superior a cero y agregar el valor de tal variable en el cuadro.

La Oficina del Censo de los Estados Unidos ha hecho del CENTS un programa de dominio público. No coloca ninguna restricción para el empleo de cualquier parte del programa. Por el contrario, con todo gusto proporciona a los posibles usuarios, sobre todo a las oficinas de estadística de los países en desarrollo, toda la documentación existente, incluso ejemplares del programa.